DOCUMENT RESUME

ED 163 025 TM 007 810

AUTHOR Karweit, Nancy L.; And Others
TITLE Examining the Credibility of Offsetting Contextual Effects. Report No. 250.
INSTITUTION Johns Hopkins Univ., Baltimore, Md. Center for Social Organization of Schools.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE May 78
CONTRACT NIE-G-78-0110
NOTE 39p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Analysis of Covariance; *Data Analysis; *Educational Research; *High School Role; Institutional Characteristics; Models; Multiple Regression Analysis; *Research Design; *Research Methodology; Research Problems; Senior High Schools
IDENTIFIERS *Multicollinearity

ABSTRACT

A strategy is proposed for dealing with educational research problems, using as an example a study based on data that had already been collected and using that data to discover how differences in school characteristics may result in differences in their students. The research method is based on concepts of analysis of covariance and is concerned with three levels of effects: (1) compositional effects due to initial differences in input variables; (2) contextual effects that have the same effect on all members of a single group; and (3) structural effects that correspond to unequal slopes in regression lines. Three relevant theoretical perspectives are: the small group perspective, interpersonal attraction theories, and the ecological or resource distribution tradition. Interaction effects may provide real clues about actual school differences. The research should begin with a comprehensive structure, and at least two plausible versions of that structure should be proposed with significantly different parameter values. Then the various proposed structures should be compared against the available data and a sensitivity analysis should be conducted to determine whether the data support any one of the structures. (CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED163025

Report No. 250

May 1978

EXAMINING THE CREDIBILITY OF OFFSETTING CONTEXTUAL EFFECTS

Nancy L. Karweit, James Fennessey and Denise C. Daiger

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Nancy Karweit

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND TO THE ERIC SYSTEM."

The Johns Hopkins University

STAFF

Edward L. McDill, Co-director

James M. McPartland, Co-director

Karl L. Alexander

Henry J. Becker

Bernard L. Blackburn

Vicky C. Brown

Martha A. Cook

Denise C. Daiger

Joyce L. Epstein

James J. Fennessey

Gary D. Gottfredson

Linda S. Gottfredson

Larry J. Griffin

Edward J. Harsch

John H. Hollifield

Lawrence F. Howe

Barbara J. Hucksoll

Nancy L. Karweit

Hazel G. Kennedy

Marshall B. Leavey

Anne McLaren

James M. Richards, Jr.

Margaret Ann Ricks

Ricahard R. Scott

Robert E. Slavin

Charles B. Thomas

Gail E. Thomas

EXAMINING THE CREDIBILITY OF OFFSETTING CONTEXTUAL EFFECTS

Contract No. NIE-G-78-0110

Nancy L. Karweit
James Fennessey
Denise C. Daiger

Report No. 250

May 1978

Published by the Center for Social Organization of Schools, supported in part as a research and development center by funds from the United States National Institute of Education, Department of Health, Education and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the Institute should be inferred.

The Johns Hopkins University
Baltimore, Maryland

Introductory Statement

The Center for Social Organization of Schools has two primary objectives: to develop a scientific knowledge of how schools affect their students, and to use this knowledge to develop better school practices and organization.

The Center works through three programs to achieve its objectives. The Policy Studies in School Desegregation program applies the basic theories of social organization of schools to study the internal conditions of desegregated schools, the feasibility of alternative desegregation policies, and the interrelation of school desegregation with other equity issues such as housing and job desegregation. The School Organization program is currently concerned with authority-control structures, task structures, reward systems, and peer group processes in schools. It has produced a large-scale study of the effects of open schools, has developed the Teams-Games-Tournament (TGT) instructional process for teaching various subjects in elementary and secondary schools, and has produced a computerized system for school-wide attendance monitoring. The School Process and Career Development program is studying transitions from high school to post secondary institutions and the role of schooling in the development of career plans and the actualization of labor market outcomes.

This report, prepared by the School Organization Program, explores the improvement of methodology for examining how schools affect students.

Abstract

This paper explores a strategy for conducting educational research with limited data and sometimes inadequate methodology, using the study of contextual effects as an example. Essentially, the strategy requires defining the compositional, contextual, and structural effects; elaborating the theoretical perspectives that could explain how influence is exerted, and investigating interaction effects for indications of real school differences. The research should begin with a comprehensive structure; at least two versions of that structure should be proposed, with different parameter values; the proposed structures should be compared against the available data, and finally, a sensitivity analysis should be conducted.

1.0 Introduction

How schools affect the students who attend them has been a subject of much research and policy concern in the last decade. Despite wide interest in the topic, research results have not provided consistent knowledge about how differences among schools influence students. Reacting to the limited research findings, commentators often note particular shortcomings of the present methodology or limitations of the data used. A fairly standard list of methodological problems--multicollinearity, measurement error, inappropriate (cross-sectional) data, inappropriate units of analysis, non-independence of sampling units--is now routinely itemized as a partial explanation for the paucity of findings. The implication of this explanation is that, in the future, the methodological situation will be better.

However, rather than looking optimistically to some future day when these problems will have disappeared, this paper advocates a different view. We suggest that future studies will still face these same methodological problems, and we attempt, using examples of contextual effects studies, to formulate a preliminary version of a research strategy cognizant of these circumstances. In short, this paper suggests a pragmatic response to the realities of conducting research with limited data and a sometimes inappropriate methodology. This strategy builds on the usual analysis of covariance logic.

1.1 Overview of Contextual Effects Research in Schools

Researchers have had a longstanding and active interest in "contextual" effects in schools, but contextual effect has not usually been precisely

defined. Typically, it pertains to the effect of being a member of a particular group, or in this case, being a student at a particular school. Basically these studies have attempted to show that the SES composition or ability composition of the school will influence the student's educational orientations and achievements. More recently, researchers have documented the existence of multiple contexts, offsetting in nature, which operate to mask conceptually important, although numerically small, contextual effects (Alexander and Eckland, 1975; Nelson, 1972; Johnson, 1971). For students of a given ability level, attending school with high ability students depresses educational expectations while attending school with high SES students appears to enhance expectations (e.g. Alexander and Eckland, 1975). Because high ability composition and high SES composition usually occur together, these offsetting effects may not be apparent if only the SES composition of the school is considered. Consequently, to understand how school composition influences an individual student's expectations, both SES and ability contexts have to be considered, but as separate influences.

Beyond the particular complication of the offsetting effects, future studies of the schooling process need to deal with several theoretical issues which, of course, affect the methodological issues. Three particular theoretical topics--definition, elaboration, and interaction--are briefly discussed below.

2.0 Some Theoretical Problems

2.1 Definition

Multi-level effects have historically been imprecisely defined.

Borrowing appropriate terminology from the analysis of covariance framework, we distinguish three types of multi-level effects: (1) compositional; (2) contextual; and (3) structural.

A compositional effect is not the result of the operation of a within-school process, or even of a group-level process. Rather, it is simply the necessary logical consequence of within-group relationships and the fact that different groups have differing average values on the input variables. This sort of effect is adjusted out of the net group average by the usual composition adjustor in analysis of covariance. In many situations, these composition effects are of substantive interest, and are sometimes referred to as demographic effects or ecological effects.

A contextual effect describes some characteristic of the group which changes the level of an outcome variable by the same amount for all persons in the same group. It corresponds to a net main effect in traditional analysis of covariance. However, we use the term contextual effect only when there is clear theoretical as well as statistical support for the interpretation. In short, contextual effects are not identical with net group differences, but they are a reasonable possibility. Also, in our usage, a contextual effect may be either positive or negative in direction, and is strictly a formal term. That is, the "context" believed to be exerting the effect may be defined by any theoretically acceptable group-level property, aggregative or global.

The third type of effect to be considered is a structural effect. The term structure refers to the structural equations of the model,

with numerically specified coefficients. A structural effect is the impact of a (relatively) macro situation on the parameters of a nested, micro-level process. In other words, the coefficients of the micro process differ depending upon the characteristics of the macro situation. This type of effect corresponds to the traditional unequal slopes situation in analysis of covariance.

2.2 Elaboration

There are several theoretical perspectives that elaborate how the setting in which an individual is located can exert influence. We list three: (1) the small groups perspective; (2) interpersonal attraction theories; and (3) the ecological or resource distribution tradition. Despite the extent of these possibilities, most contextual effects studies have not employed existing theories carefully, nor have they developed a separate theory of influence in schools. However, elaboration of the mechanisms of influence is critical. As one simple example, consider alternate theories of peer influence and their corresponding measurement implications. One prevalent theory of peer influence would argue that the similarity of peers, by an interpersonal reinforcing mechanism, serves to enhance values and norms. Another theory (relative deprivation) suggests that it is the difference of an individual from the group norms, not the similarity, which affects behavior. In one situation, we would predict a strong positive effect if $x_i - \bar{x}$ were small, in the other a strong positive effect if $x_i - \bar{x}$ were large. Either result could be plausible depending upon the type of influence posited.

2.3 Interaction

The existence of an interaction effect, as evidenced by the unequal slopes situation in analysis of covariance, is frequently encountered. When this occurs, it is most often viewed as a nuisance. As suggested by Burstein (e.g. 1977) and also by Karweit and Fennessey (1977), such interactions may provide real clues about how schools actually differ. Rather than regarding the interaction as a nuisance, we see it as an opportunity for documenting a schooling effect. In terms of the contextual effects studies, the problem becomes one of establishing reasonable guidelines for determining when variation among the "structural" parameters (the slopes) can be regarded as reflecting actual differences in process, not differential measurement error or instabilities in the estimates of the slopes. Another related issue is whether the process works the same for all students within the same school. Instead of a schoolwide contextual effect, it may be more reasonable to propose the existence of sex- or curriculum-specific contexts within which outcomes are determined.

3.0 Contextual Analysis: An Example

We now present results from a "contextual" analysis of 11,729 students in forty-eight non-southern, metropolitan four-year high schools. These schools constitute a subsample from the Equality of Educational Opportunity (EEOR) data (Coleman, et al., 1966), the same subsample used by Heyns in her 1974 analysis of selection and stratification within schools. These data were selected for replication of previous contextual analyses primarily because they overrepresent

non-white students and include schools which are more likely than the average American high school to be racially integrated. Prior research on the effects of school context on individuals has been done on samples which underrepresented large, racially diverse urban high schools.

The model to be fitted to these data is in Figure 1. It is similar to the traditional model employed by previous contextual effects researchers (Alexander and Eckland, 1975) except that it omits peer influences. This specification was necessitated by the lack of peer information in the EEOR data.

Figure 1 About Here

Table 1 shows the estimated reduced form equations of the model. SEScon has several positive effects on EDEX, a direct effect of .166 and indirect effects through teacher expectation and parental expectation. However, we must interpret the TCHEX and PAREX coefficients cautiously: being based on students' reports, they are likely to have correlated measurement error. APTcon has a negative effect on G.P.A., PAREX and EDEX. Also, APTcon has a positive effect on curriculum enrollment. The latter finding is contrary to prior research (Alexander and Eckland, 1975; Alwin, 1976) which reported a negative effect of aptitude context on curriculum enrollments.

Tables 1 and 2 About Here

Fitting the structural form of the model (Table 2), shows that aptitude context tends indirectly to lower educational expectations by lowering G.P.A. and parents' expectations. As noted earlier, APTcon

indirectly raises educational expectations by raising one's chance of enrollment in a college preparatory curriculum. However, these structural variables do not account for the total effect of the aptitude context. A more adequate specification of the model would include observed measures of more proximate influences, such as the degree of academic competition in the school.

Similarly, this analysis leaves us with many questions concerning the relationship between SES context and educational expectations. SEScon indirectly raises significant adults' expectations, but the model leaves unexplained a sizeable "direct" effect. The inclusion of an explicit measure tapping the mechanism thought to operate (whether based upon a reference group perspective, balance theory, ecological theory, or some other perspective) would account for the direct effect of SES context on educational expectations.

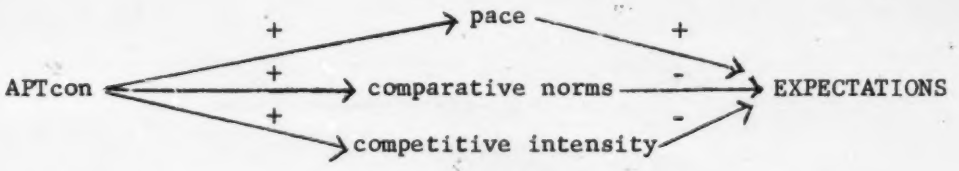
3.1 Multicollinearity Analyses

Before accepting these findings as evidence of offsetting contextual effects, we should, however, examine the possibility that these effects are artifactual. Because the two context measures are typically highly correlated (.801 in these data), we must consider the possibility that the offsetting effects arose as a result of the "tipping effect" well documented by Gordon (1968). He showed that when there is high correlation among regressors, the coefficients become highly sensitive to minor disturbances in the correlation matrix. This artifact often results in large regression weights of opposite sign.

Although multicollinearity has been recognized as a problem in contextual analysis (Alwin, 1976), it has never received a thorough substantive treatment in this literature. Roughly, multicollinearity is an overlap among the regressors. But, a neglected though important question concerns the degree to which such overlap is viewed as "accidental," or as necessary according to the model. Accidental multicollinearity is that which is in no way required by the theoretical model. Necessary multicollinearity, on the other hand, is logically implied by some aspect of the specification imposed. The literature on status inconsistency provides an example of necessary collinearity where original status, destination status, and distance moved are all regarded as theoretically relevant variables in a linear model (Blalock, 1967). Yet, these three variables are necessarily and totally collinear. When such complete multicollinearity arises from a logical specification, we usually recognize it immediately and deal with it by redefinition. On the other hand, when we consider the last equation in a complex structural model, many of its regressors have an implied degree of multicollinearity simply as a result of the previous equations. This fact, obvious enough when stated, is seldom taken into account in the statistical literature on dealing with multicollinearity. Given observed multicollinearity among some regressors which are themselves endogenous, we must recognize that any attempts to deal with it will imply changes in the specification and parameter values generating the collinear variables.

This distinction between accidental and necessary multicollinearity depends, of course, upon the degree of detail in the model as it is

operationally used. At the abstract level, the model we are considering may include several distinct processes, but the operationalization may make these processes necessarily collinear (cf. Firebaugh, 1977). For example, in one abstract model the aptitude context may tend to raise expectations by enabling instruction to proceed more rapidly, thus raising achievement as a byproduct, and leading to study of more advanced topics; similarly, the aptitude context may tend to cause informal norms of achievement to be higher, in the comparative sense. Also, high aptitude may intensify the competitive pattern of activity. The high comparison norms and the greater intensity of competition may tend to keep achievement high, but at the same time produce feelings of inadequacy in students whose scores are thus, relatively, not so high, and whose work is thus, relatively, less often rewarded. Diagrammatically, this might be shown as:



If this model is viewed as the "true" one, but only APTcon and EXPECTATIONS are measured, there is no operational way to disentangle the processes. We could obtain only a composite measure, showing the net resultant of all three linkages. Another difficulty arises, however, if we now introduce SEScon as well into the discussion. To the extent that SEScon is correlated with the portion of APTcon causing pacing differences, we will have the following consequence: the effect of

SEScon will be overestimated, because the portion of the outcome variance due to pacing differences now goes to it. Thus, the example illustrates the kind of problem we face. Under this set of assumptions the effects of pacing--which our model links to APTcon rather than SEScon--will nevertheless turn up as a contributor to the SEScon effect. Both estimated coefficients will be biased and in a way that is consistent with previous findings. The APTcon coefficient no longer reflects the average of all three processes; the SEScon coefficient overstates the impact of SEScon.

What response is available to us for dealing with these consequences of logical multicollinearity? One approach is simply to obtain measurements on the various intervening variables. This is a direct and often reasonable strategy when research is being planned. It is of little value, however, to someone working with an existing data set collected and computerized at great expense. For that person, the analogous strategy is to carry out some sensitivity studies. First, the researcher should specify some likely alternative patterns among these variables, observed and unobserved. Next he should examine how different assumed patterns affect the main inferences about multi-level processes. In short, when data are lacking, we may still introduce data-equivalent information, based on parallel studies and/or expert judgment.

Presuming that we have dealt as well as possible with this necessary collinearity, be it total or partial, we are left with the likely existence of accidental collinearity as well. In this paper we are not addressing numberical accuracy issues (Longley, 1967).

Various ad hoc measures have been proposed to detect multicollinearity. However, recent work (cf. Rockwell, 1975; Mason, Gunst, and Webster, 1975; Silvey, 1969) has made clear that three descriptive indexes are important. All of these are derived from R_{XX}, the matrix of correlations among the regressors. The first is simply the determinant of R_{XX}. If it is "sufficiently" large, then no serious collinearities are present. The second set of numbers is the diagonal elements of the inverse of R_{XX}. These diagonals identify the regressors which are most involved in a multicollinearity (cf. Farrar and Glauber, 1967). The third set of numbers is the eigenvalues of R_{XX}. These show, by the size of the smallest eigenvalues, how much collinearity is present and of what dimension it is. Unfortunately, for none of these indexes is there a satisfactory set of cutpoints. We know that small determinants and small eigenvalues are generally indicative of collinearity, but we do not know how small is "small." This is a topic on which simulation work may be quite helpful.

The so-called Haitovsky heuristic test (Haitovsky, 1969; Rockwell, 1975) of the determinant is in fact of little direct value to an investigator. It depends upon the number of cases in the sample, but most of the unfortunate consequences of collinearity do not. Also, it tests an irrelevant hypothesis--namely that the observed matrix may have been drawn from a population in which the corresponding matrix is singular. If the test is failed, it is a sure sign of trouble. However, if it is passed, that is no assurance that the problem of collinearity will be avoided.

The eigenvalue approach also has its drawbacks. Most importantly, it yields a set of numbers rather than a single number, so there is still some problem of reducing these to a single value. Although there are several summary numbers in the literature (condition number, Hoerl & Kennard's 1970 index), none gives more than ordinal information.

What are the consequences of high accidental collinearity? The consequences are not just in terms of sampling fluctuations, although there appears to be some confusion on this point. Strictly, the consequence is that the inferences about B are highly sensitive to any change in R_{XX} or R_{YY}. That is, even with a perfect sample, the coefficients would still be highly dependent on the exact data pattern of X and Y. They would not generalize even to a slightly different population. Stating this differently, if we observed the complete population of interest, then our coefficients would logically have zero sampling variability. They would necessarily be equal to the population value. Yet, their weakness is that the results from that particular population would be very poor guides to what might be expected from a slightly different population.

The final question is, what can we do about the multicollinearity issue? Various strategies have been proposed for dealing with collinearity. They include variable selection methods such as stepwise regression. But for research examining a structural model, these techniques are simply irrelevant. Another class of techniques known as biased regression offers some promise. Ridge regression is the best known example of this class of techniques. Recent work by statisticians

(cf. Hoerl, Kennard, Baldwin, 1975; Dempster, Schatzoff, Wermuth, 1977) indicates that ridge regression may be a highly useful procedure for dealing with collinearity of the accidental sort.

Ridge regression is performed by adding a small constant (.01, .02, etc.) to the diagonal of R_{XX} and then doing the regression as usual. The advantage of the technique is that it accepts some small bias in the estimated coefficients in return for a large reduction in their variance, thus gaining an overall improvement in mean squared error of the coefficients (cf. Winer, 1978). Two limitations should be kept in mind. One, the technique is useful only when collinearity is severe. Two, it requires some discretionary judgment by the researcher as to the choice of the constant to be added.

For our data, the coefficients computed with ridge regression, using several reasonable choices of the ridge constant, did not differ noticeably from those obtained with ordinary regression. It occurred to us, however, that the observed <u>difference</u> between the regression weights for the context measures might not be significantly different from zero, though each coefficient individually achieved significance. Such a situation might arise because the regression weights are highly related and their covariance is not taken into account when their difference is examined. The appropriate test for these differences uses a denominator calculated as $SQRT(Var[B_i] + Var[B_j] - 2 * cov\ [B_iB_j])$ However, the offsetting effects found in our study retained a significant difference even when put to this more proper test. Thus, this particular artifact was ruled out as an explanation of these differences.

3.2 Race and Sex Subgroup Analyses

Given the enhanced credibility of these offsetting contextual effects, we wanted to determine whether the effects would hold up for different subgroups of our sample. A test for equality of regression slopes showed that the processes for the four major race/sex subgroups do indeed differ beyond what would be expected by chance. Therefore, separate analyses on each subgroup were performed to provide descriptive information about the distinct process.

Tables 3-10 About Here

The present model does not explain educational expectations for blacks as well as it does for whites (R^2 = .652 for white females; R^2 = .434 for black males). The contextual effects are never offsetting--even in sign--for the blacks. In fact, the direct effect of APT_{con} on EDEX is never significant for blacks. Among blacks, SES context has different influences on EDEX for males and females (males b = .072 n.s.; females b = -.203, $p \leq .05$). These results could be accounted for by differential measurement error (Bielby, Hauser and Featherman, 1977). However, before adopting this conclusion, more explicit information on how schools affect individuals is required. Perhaps the opportunity structure in high status schools is such that black females have less access to those resources which enhance educational expectations, or the structure of peer associations may isolate them from college-oriented peers.

The white subgroup analyses likewise point up the need for more precise theoretical bases. Aptitude context lowers G.P.A. and parental

expe[illegible]les, thus indirectly lowering educational e[illegible]itude context has no effect on G.P.A. for whi[illegible] specified does not help explain these sex differences. A direct measure of perceived competition might clarify this situation. Perhaps males respond more to higher levels of competition than do females.

The status context variable operates in a more straightforward fashion for white females than for white males. For females, the total effect of this variable on educational expectations is indirect, through parents' expectations. For males, however, a direct positive effect remains after controlling for intervening variables. Also, while status context raises educational expectations indirectly by raising parents' expectations, it indirectly lowers expectations by lowering one's chances of being enrolled in a college curriculum.

4.0 Some Next Steps

The analysis by subgroup supplies some hints for the appropriate specification of a contextual model. It also indicates convincingly that the process by which enrollment in a particular school affects an individual's educational expectations is not uniform for all individuals. If we hope to better understand how contextual effects operate, we must then consider the position of the individual within his social setting and its social psychological consequences. Undoubtedly, such considerations present numerous and severe measurement and operational complexities. For example, operational definitions become more of a concern. Nonetheless, these methodological concerns must be recognized and dealt with simultaneously.

One unfortunate consequence of posing a more complex model is the natural tendency toward "selective attention." By this we mean that in the beginning of a study, the researcher usually outlines a wide range of relevant features of the phenomenon under study. Then, however, his attention becomes concentrated on one narrow, selected portion of the phenomenon, without corresponding attention to the conditional nature of the results being studied. Moreover, assumptions made about the incidental features of the situation frequently tend to maximize desired results or interpretations. This selective attention problem--looking where the light is best rather than where the coin may have fallen--occurs between studies as well as within studies. Such behavior is convenient and tempting, but its attractiveness is shortlived. It has created a hiatus between research results and substantive complexity, which has prevented the accumulation of credible research findings.

Clearly,concentration of energy and narrowed focus is unavoidable and even desirable as an aspect of research strategy. The peril of selective attention is that the particular conditions under which the results were obtained are forgotten and that logically inconsistent assumptions are not exposed. Given the likelihood of these conditions, the use of formally-represented, relatively comprehensive models, is the only reliable way to proceed. Even though the particular project may focus on a small fraction of the many aspects of a problem, it must be able to locate itself unambiguously with respect to all.

Locating one's research with respect to prior relevant research implies the need for a comprehensive model which envelops all relevant

variables, regardless of the capacities of one's data source. This strategy demands that models should be theory-driven rather than data-driven. In developing such a comprehensive model, the researcher is forced to choose specific numerical values for all the necessary parameters. At least two alternative concrete versions of the model (hereafter referred to as structures) should be developed, which will be formally compatible but distinct. Acceptable assumptions about incidental parameters will thus be explicit and competing viewpoints held concerning the pattern of relationships expected will be reflected.

The LISREL program and its companion programs SIFASP and COFAMM, developed recently by Joreskog and his associates (Joreskog, 1973; Joreskog and Sorbom, 1976; Sorbom and Joreskog, 1976), provide a highly convenient general tool for representing and exploring these parameter configurations. The programs effectively implement a distinctive paradigm of research inference, one which embodies flexibly and formally the notions of substantive model and measurement model. The models are technically restricted to the class of linear models, but this is a limitation that can be tolerated quite well.

The actual choice of the parameter values which defines a particular structure is perhaps the most difficult part of the entire strategy. Several possibilities can be mentioned. In some instances, the parameter value needed can be drawn directly from previous studies. In other cases, there will be no directly relevant quantitative evidence, and the investigator will have to rely upon expert judgment, either his own or that of some recognized experts. Every effort should be

made to identify those parameters which can be accepted consensually, and to improve estimating approaches for those which are controversial or unknown.

After at least two such concrete structures, and usually several, have been specified, the main data analysis is carried out. It consists of comparing the several proposed structures against the dataset at hand. The comparison has four parts. One is concerned with the absolute closeness of fit between the structures and the observed data, and also with the relative closeness--are some of the structures considerably closer to the data than others? In addition, the closeness is to be assessed globally and locally. That is, an overall measure of goodness of fit is needed (perhaps based on the chi-square statistic provided by LISREL), but interest also is devoted to the particular ways in which the data departs from each structure (perhaps using indices of the residual covariances or the derivatives of the fixed parameters). Thus, in this strategy the aim is not merely to show, as is often done, that the data are consistent with one favorite structure. Nor is it to show that the data are inconsistent with some ridiculously simple null hypothesis. Such comparisons fall short of real scientific utility. Instead, the data analysis is a detailed comparison of several plausible alternatives (cf. Fennessey, 1977).

The final stage in the overall strategy is the execution of an intelligent sensitivity analysis (Land and Felson, 1978). Here the aim is to ascertain the degree to which model fit, and specific predicted outcomes, are dependent upon relatively minor changes in certain parameter settings, or specific characteristics of the input variables.

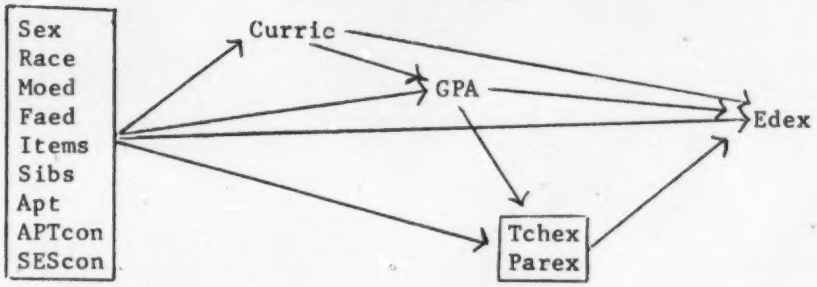
In short, this step examines the stability of the results against various kinds of perturbations in the real circumstances or the research design.

Quite often, the sensitivity analysis will also involve asking whether the processes (substantive and methodological) operate homogeneously for various subpopulations. The subpopulations may be defined in terms of individual student characteristics (e.g., sex, race, age, ability, etc.), or in terms of relations between the individuals and the social systems (e.g., membership in School A versus school B, college versus vocational track, or clique A versus B, etc.)

The strategy as just outlined, although complex, is feasible and realistic. Moreover, the short-run benefits obtainable from more simplistic approaches appear to have become exhausted. Researchers, clients, and sponsors alike are becoming impatient for a better approach. This strategy, or one along similar lines, will yield more enduring benefits because it represents a more realistic balancing between knowledge and ignorance, and between uncertainty and ideology.

Figure 1

Contextual Effects Model[1/]



Blocked variables enter the model simultaneously. Variable abbreviations are as follows:

Sex: male coded 1, female coded 2
Race: white coded 1, other coded 0
Moed: mother's education
Faed: father's education
Items: household acquisitions
Sibs: number of brothers and sisters
Apt: verbal ability
APTcon: school average verbal ability
SEScon: school average father's education
Curric: enrollment in college preparatory program
GPA: grade point average
Tchex: encouragement from teacher or counselor to continue schooling after high school
Parex: amount of education parents desire for student
Edex: amount of education student desires

[1/]All variables except APT are coded from students' self-reports.

Table 1

	Reduced Form				Total Sample N = 11729					
	Curric		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo.ed.	.031[a]	.003[b]	.019	.005	.009	.003	.094	.007	.091	.008
Fa.ed.	.026	.002	.025	.004	.012	.002	.084	.006	.085	.007
Items	.256	.036	(-.010)	.056	.253	.035	1.426	.091	1.2	.102
Sibs	-.011	.002	-.019	.003	-.006	.002	-.034	.005	-.029	.006
Apt	.015	.000	.021	.000	.006	.000	.025	.001	.036	.001
Sex	-.052	.008	.154	.012	-.027	.007	-.520	.091	-.458	.021
Race	-.127	.012	(-.029)	.018	-.084	.011	-.448	.029	-.568	.033
SEScon	(-.014)	.009	(-.013)	.014	.018	.009	.151	.022	.166	.025
APTcon	.006	.001	-.004	.002	(-.001)	.001	-.011	.003	-.009	.003
R^2	.323		.221		.079		.306		.317	

(Coefficient) does not reach twice its standard error

a: metric coefficient
b: standard error

Table 2

	Full Form				Total Sample N = 11729					
	Curric		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo.ed.	.031[a]	.003[b]	.012	.005	(.004)	.003	.064	.007	.012	.006
Fa.ed.	.026	.002	.019	.004	.007	.002	.057	.006	.014	.005
Items	.256	.036	(-.069)	.055	.219	.034	1.199	.084	(.147)	.076
Sibs	-.011	.002	-.016	.003	(-.004)	.002	-.021	.005	(.001)	.004
Apt	.015	.000	.017	.000	.003	.000	.009	.001	.008	.001
Sex	-.052	.008	.166	.011	-.029	.007	-.496	.018	-.129	.016
Race	-.127	.012	(-.000)	.018	-.065	.011	-.330	.027	-.204	.024
SEScon	(-.014)	.009	(-.010)	.014	.020	.008	.165	.021	.081	.019
APTcon	.006	.001	-.006	.002	(-.001)	.001	-.015	.003	-.005	.002
Curric			.229	.014	.137	.009	.895	.022	.592	.021
G.P.A.					.058	.006	.142	.014	.136	.013
Teacher ex.									.153	.020
Parent ex.									.606	.008
R^2	.323		.238		.109		.407		.629	

(Coefficient) does not reach twice its standard error
a: metric coefficient
b: standard error

Table 3

Reduced Form — White Males, N = 4984

	Curric.		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
	[a]	[b]								
Mo.ed.	.021	.005	.015	.007	(.005)	.004	.067	.011	.064	.013
Fa.ed.	.023	.004	.028	.006	.009	.003	.074	.009	.084	.010
Items	.310	.069	-.231	.110	.441	.064	1.259	.165	1.096	.190
Sibs	-.012	.003	-.019	.005	(.005)	.003	-.042	.008	-.035	.009
Apt	.014	.001	.021	.001	.004	.000	.024	.001	.036	.001
SEScon	-.028	.014	(-.042)	.023	(.005)	.013	(.053)	.034	(.134)	.039
APTcon	.009	.002	(.004)	.003	(.002)	.002	.012	.004	(.005)	.005
R^2	.285		.203		.056		.256		.294	

(Coefficient) does not reach twice its standard error
a: metric coefficient
b: standard error

Table 4

Full Form — White Males N = 4984

	Curric.		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo.ed.	.021[a]	.005[b]	(.009)	.007	(.001)	.004	.047	.010	(.007)	.009
Fa.ed.	.023	.004	-.021	.006	.005	.003	.051	.008	.019	.008
Items	.310	.069	-.318	.109	.418	.063	1.051	.153	(.105)	.141
Sibs	-.012	.003	-.015	.005	(-.002)	.003	-.029	.008	(.002)	.007
Apt.	.014	.001	.017	.001	.001	.000	.009	.001	.008	.001
SEScon	-.028	.014	(-.034)	.022	(.011)	.013	.084	.031	.126	.028
APTcon	.009	.002	(.001)	.003	(.000)	.002	(.004)	.004	-.009	.003
Curric.			.283	.022	.116	.013	.807	.032	.617	.031
G.P.A.					.053	.008	.178	.020	-.172	.018
Teacher ex.									.159	.032
Parent ex.									.611	.013
R^2	.285		.228		.082		.362		.624	

(Coefficient) does not reach twice its standard error.

a: metric coefficient

b: standard error

Table 5

Reduced Form — Black Females, N = 1063

	Curric.		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo.ed.	.029[a]	.009[b]	(.025)	.014	(.012)	.009	.122	.023	.145	.025
Fa.ed.	(.015)	.008	.024	.012	(.005)	.008	(.036)	.021	(.041)	.023
Items	.339	.081	.273	.120	.186	.078	1.85	.205	1.783	.225
Sibs	(.006)	.005	.01)	.008	(-.004)	.005	(0)	.015	(.008)	.014
Apt.	.015	.001	.017	.001	.006	.001	.022	.002	.032	.003
SEScon	(-.083)	.043	(.027)	.064	(.001)	.042	(-.126)	.109	-.319	.120
APTcon	(.004)	.004	-.016	.005	(-.005)	.003	(.004)	.009	(.017)	.010
R^2	.294		.170		.063		.243		.290	

(Coefficient) does not reach twice its standard error.
a: metric coefficient
b: standard error

Table 6

Full Form — Black Females, N = 1063

	Curric.		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo.ed.	.029[a]	.009[b]	(.021)	.014	(.007)	.009	.101	.022	.056	.020
Fa.ed.	(.015)	.008	(.022)	.012	(.002)	.008	(.025)	.020	(.009)	.018
Items	.339	.081	(.224)	.121	(.133)	.078	1.606	.199	.5	.184
Sibs	(.006)	.005	(-.011)	.008	(-.005)	.005	(-.004)	.012	(.007)	.011
Apt.	.015	.001	.015	.002	.003	.001	.011	.003	.008	.002
SEScon	(-.083)	.043	(.039)	.046	(.011)	.041	(-.066)	.105	-.203	.094
APTcon	(.004)	.004	(-.016)	.064	(-.005)	.003	(.000)	.009	(.015)	.008
Curric.			.145	.005	.129	.030	.722	.076	.609	.071
G.P.A.					(.034)	.020	(-.004)	.051	.114	.046
Teacher ex									.181	.071
Parent ex									.547	.028
	.294		.178		.083		.304		.563	

(Coefficient) does not reach twice its standard error.
a: metric coefficient
b: standard error

Table 7

Reduced Form — Black Males — N = 808

	Curric.		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo.ed.	.026[a]	.011[b]	.033	.017	(.010)	.010	.082	.027	(.041)	.033
Fa.ed.	.044	.009	(.025)	.015	.019	.009	.084	.024	.084	.029
Items	.187	.082	(-.149)	.130	(.013)	.080	1.745	.206	1.676	.250
Sibs	(-.004)	.006	(.001)	.010	(.000)	.006	(-.002)	.016	(-.008)	.019
Apt	(-.013)	.001	(.010)	.002	(.004)	.001	.017	.003	.624	.003
SEScon	(-.020)	.046	(.082)	.072	(.034)	.045	(.100)	.115	(.130)	.140
Aptcon	(.003)	.004	-.011	.007	(-.006)	.004	(-.013)	.010	(-.006)	.013
R^2	.288		.077				.225		.184	

(Coefficient) does not reach twice its standard error
a: metric coefficient
b: standard error

Table 8

Full Form — Black Males — N = 808

	Curric.		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo.ed.	.026[a]	.011[b]	(.030)	.017	(.004)	.010	.062	.026	(-.024)	.027
Fa.ed.	.044	.009	(.019)	.015	(.012)	.009	.055	.023	(.007)	.024
Items	.187	.082	(-.172)	.130	(.004)	.079	1.65	.199	.618	.219
Sibs	(-.004)	.006	(.001)	.010	(.001)	.006	(.0)	.015	(-.005)	.016
Apt	.013	.001	.009	.002	(.002)	.001	.008	.003	(.005)	.003
Sescon	(-.020)	.046	(.084)	.072	(.029)	.044	(.102)	.111	(.072)	.117
Aptcon	(.003)	.004	(-.011)	.006	(-.006)	.004	(-.013)	.010	(.001)	.011
Curric.			.125	.055	.114	.034	.608	.086	.576	.093
G.P.A.					.086	.022	.128	.054	(.187)	.058
Teacher ex.									.202	.095
Parent ex.									.550	.038
R^2	.288		.083		.076		.279		.434	

(Coefficient) does not reach twice its standard error
a: metric coefficient
b: standard error

Table 9

Reduced Form — White Females, N = 4874

	Curric.		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo.ed.	.042[a]	.005[b]	.019	.007	.009	.005	.122	.012	.120	.013
Fa.ed.	.031	.004	.022	.006	.017	.004	.105	.010	.098	.011
Items	(.112)	.068	(.031)	.099	.271	.067	.930	.174	(.352)	.190
Sibs	-.020	.003	-.027	.005	-.009	.003	-.054	.009	-.052	.009
Apt .	.016	.000	.023	.001	.008	.000	.027	.001	.037	.001
SEScon	(-.010)	.015	(-.027)	.021	(.008)	.014	.209	.037	.156	.041
APTcon	(.004)	.002	-.008	.003	(.002)	.002	-.027	.005	-.014	.006
R^2	.320		.263		.120		.297		.353	

(Coefficient) does not reach twice its standard error.
a: metric coefficient
b: standard error

Table 10

Full Form — White Females N = 4874

	Curric.		G.P.A.		Teacher ex.		Parent ex.		Ed. ex.	
Mo. ed.	.042[a]	.005[b]	(.01)	.007	(.002)	.005	.076	.011	(.018)	.01
Fa. ed.	.031	.004	.015	.006	.011	.004	.070	.009	(.011)	.008
Items	(.112)	.068	(.008)	.098	.251	.066	.809	.158	-.331	.137
Sibs	-.020	.003	-.023	.005	(-.004)	.003	-.030	.008	(-.003)	.007
Apt	.016	.000	.02	.001	.004	.001	.008	.001	.008	.001
SEScon	(-.010)	.015	(-.025)	.021	(.011)	.014	.223	.034	(.033)	.029
APTcon	(.004)	.002	-.009	.003	(.002)	.002	-.030	.005	(.001)	.004
Curric			.203	.021	.157	.014	1.037	.034	.557	.032
G.P.A.					.063	.010	.125	.023	.101	.020
Teacher ex.									.146	.030
Parent ex.									.621	.012
R^2	.353		.278		.153		.422		.652	

(Coefficient) does not reach twice its standard error.

a: metric coefficient
b: standard error

References

Alexander, K. and B. Eckland
1975 Contextual effects in the high school educational attainment process. American Sociological Review, 40 (June): 402-416.

Alwin, Duane
1976 Structural equation models of contextual effects. Paper presented at Conference on Methodology for Aggregating Data in Educational Research, Stanford University.

Bielby, William T., Robert M. Hauser, and David L. Featherman
1977 Response errors of black and nonblack males in models of the integenerational transmission of socioeconomic status. American Journal of Sociology, 82: 1242-1288.

Blalock, Hubert M. Jr.
1967 Status inconsistency, social mobility, status integration, and structural effects. American Sociological Review, 32: 790-801.

Burstein, Leigh
1977 Proposal: Analysis of multilevel data in educational research and evaluation.

Coleman, James, E. Q. Campbell, C.J. Hobson, J. McPartland, A.M. Mood,
1966 F.D. Weinfeld and R. York
Equality of Educational Opportunity. Washington: United States Government Printing Office.

Dempster, Arthur P., M. Schatzoff, and N. Wermuth
1977 A simulation study of alternatives to ordinary least squares. Journal of the American Statistical Association, 72: 77-106.

Farrar, Donald E. and R. R. Glauber
1967 Multicollinearity in regression analysis: The problem revisited. Review of Economics and Statistics, 49: 92-109.

Fennessey, James
1977 Improving inference for social research and social policy: The Beyesian Paradigm. Social Science Research, 6: 309-327.

Firebaugh, Glenn
1977 Groups as contexts and frog ponds: Some neglected considerations unpublished paper. Vanderbilt University.

Gordon, R.
1968 Issues in Multiple Regression. American Journal of Sociology, 73: 592-616.

Haitovsky, Yoel
1969 Multicollinearity in regression analysis: Comment. Review of Economics and Statistics, 59: 486-489.

Heyns, Barbara
1974 Social selection and stratification within schools. American Journal of Sociology, 79: 1434-1451.

Hoerl, Arthur E. and R. W. Kennard
1970 Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12: 55-67.

Hoerl, A.E., R.W. Kennard and K.F. Baldwin
1975 Ridge Regression:Some Simulations. Communications in Statistics , 4:105-23.

Johnson, D.
1971 School impact of educational aspirations of U.S. public high school seniors: A methodological and causal analysis. Doctoral dissertation. Vanderbilt University. Ann Arbor, Michigan. University Microfilms No. 72-15, 476.

Joreskog, Karl G.
1973 A general method for estimating a linear structural equation system. In Arthur S. Goldberger and Otis Dudley Duncan, (Eds.), Structural Equation Models in the Social Sciences, New York: Seminar Press.

Joreskog, Karl G. and Dag Sorbom
1976 LISREL III: Estimation of Linear Structural Equation Systems by Maximum Likelihood Methods. National Educational Resources: Chicago.

Land, Kenneth C. and Marcus Felson
1978 Sensitivity analyses of arbitrarily identified simultaneous equation models. Sociological Methods and Research, 6: 283-307.

Longley, James
1967 An appraisal of least squares programs for the electronic computer from the point of view of the user. Journal of the American Statistical Association, 62: 819-829.

Mason, Robert L., R. F. Gunst and J. T. Webster
1975 Regression analysis and problems of multicollinearity. Communications in Statistics, 4: 277-292.

Nelson, Joel I.
1972 High school context and college plans: The impact of social structure on aspirations. American Sociological Review, 37, Number 2: 143-148.

Rockwell, Richard C.
1975 Assessment of multicollinearity: The Haitovsky test of the determinant. Sociological Methods and Research, 308-320.

Silvey , S. D.
1969 Multicollinearity and imprecise estimation. Journal of the Royal Statistical Society, Series B, 31:539-552.

Sorbom, Dag and Karl G. Joreskog
1976 COFAMM: Confirmatory Factor Analysis with Model Modification. National Educational Resources, Chicago.

Winer, B. J.
1978 Trading bias for reduced mean squared error. Annual Review of Psychology, 1978, pp. 647-681.